Fraudulent URL and Credit Card Transaction Detection System Using Machine Learning

Dr. S Geetha
Associate Professor, Dept of CSE
BNM Institute of Technology
Bangalore, India
geetha.s@bnmit.in

Yusuf Mohammed Khan
Dept of CSE
BNM Institute of Technology
Bangalore, India
ymklmfcb@gmail.com

Rohan Sujay
Dept of CSE
BNM Institute of Technology
Bangalore, India
rohansujay17@gmail.com

Sai Pavan Yoganand
Dept of CSE
BNM Institute of Technology
Bangalore, India
spavan2812@gmail.com

Rohan B
Dept of CSE
BNM Institute of Technology
Bangalore, India
devurohan@gmail.com

Abstract- **Cybersecurity encompasses operation technology, information which are closely connected practices, technologies and concepts. The use of this information offensively to attack others is the basic feature of cybersecurity. We will focus on malicious URLs and fraudulent credit card transactions as the two key cybersecurity flaws. As we all know URLs have become an important part of everyone's life. It is important to detect which URLs are malicious and which one is safe. Malicious URLs can potentially cause a huge economic loss to any country, organization or a human being. In this modern world monetary transactions are made online with the help of credit cards, despite its easy usability, security loopholes and user's carelessness lead to fraudulent transactions as hackers get access to the credit card's sensitive data, Causing an annual loss of nearly 5 billion dollars. Security tools are always in development to solve and detect malicious activity and security loopholes. In this paper, Machine Learning algorithms are used to create API's that detect security loopholes and our API tends to detect malicious URLs and fraudulent credit card transactions and host it on the web so that users don't have to download the software files locally. This paper is in correspondence to all results obtained from previous models, that gave a general accuracy of 70% to 90%**

INDEX TERMS- MALICIOUS URL, DETECTION, CREDIT CARD, FRAUDULENT, API

I. INTRODUCTION

In the modern world, where progress and breakthroughs are happening quickly every day and technology is playing a significant role in this transformation, technology is susceptible to damaging attacks by cybercriminals employing a number of faults and vulnerabilities found in the devices and the people that utilize them. A link that has been constructed with the intention of supporting scams, attacks, or frauds is known as a malicious URL.A malicious URL is one that was made specifically to spread fraud, attacks, and scams.

One is always in the risk of downloading ransomware or viruses by accidentally or diligently clicking on an infected URL which won't seem infected to the eye and they can damage your computer or, in the case of a business, your network. You could be induced to provide private information on a phony website using a malicious URL. You should be aware that there are other forms of hazards besides connections to malware that can spread online. Spam and phishing are the two scams involving harmful URLs that occur most frequently.

A credit card is a tiny, handy piece of plastic that holds personal data about the owner, like a signature or photo. Automated teller machines (ATMs), retail readers, banks, and online internet banking systems may now read the data on a credit or debit card. They must have a special card number; this is crucial. Its security is reliant on both the credit card number's confidentiality and the plastic card's physical security. A large rise in fraudulent activity has been brought on by the swift increase in credit card transaction volume. The term "credit card fraud" refers to a variety of theft and fraud schemes that employ a credit card to provide fraudulent funds for a specific transaction.

II. LITERATURE SURVEY

[1] In order to improve classification performance, they designed and created an intelligent URL detection model in [1] that uses attributes based on cyber threat data. CTI-MURLD is the model proposed, consists of three major components. The initial section is for gathering the features. Features based on URLs, features based on Whois information and characteristics based on intelligence of cyber threat, are the three categories of features retrieved. The threat elements are investigated with Google and Whois information. Then the second step is feature extraction, followed by representation and those are selected as part of data processing. Those features are represented using TF-IDF, selected using mutual information and extracted using N-grams. The third component involves making decisions and classifications. Thus, as a result, the model yielded 96.7% precision with RF, 95.69% with DT and 94.4% with the CNN based classifier.

In [2], They suggested using widely used classification approaches as decision trees, extreme learning machines, K-Nearest Neighbor (K-NN), multilayer perceptron, and support vector machines (SVM) in [2] to determine the accuracy for fraud detection. Additionally, they suggested a model in [2] that included DT, SVM, and KNN models, which considerably increased prediction accuracy. In addition to this, two web services have been implemented in [2] for efficient data sharing across heterogeneous platforms: Simple Object Access Protocol (SOAP) and Representational State Transfer (REST). The credit card fraud classification data was the source of the information used in this investigation. It has

978-1-6654-9781-7/22/$31.00 ©2022 IEEE

30000 rows and 23 columns, 690000 data points. It has been noted that the SVM model has the highest accuracy rate of all the models at 81.63%. The correctly predicative class in the suggested model has fewer false alarms overall. The observed accuracy percentage for the suggested hybrid model is 82.58%, which is greater than the prediction accuracy value for the separate models.

In [3], ML methods are constructed to detect fraud credit transactions. Different algorithms such as SVM, RF, ET, XGBoost and others were used in the study. Each technique, combined with the AdaBoost boosted the resilience. The primary contribution of [3] is a comparison of multiple ML techniques on a dataset available publicly which includes genuine transactions. DT without SMOTE gave a recall of 75.75% and with SMOTE gave 99%. Without SMOTE-AdaBoost, the DT gave a precision of 79.8% and with SMOTE gave a precision of 98.79%. The organizational editing prior to formatting enhanced the MCC from 0.78 to 0.98.

A public dataset of 45,000 URLs was used to train the model used in [4]. OpenPhish website provides harmful URLs using the best classifiers. It contrasts the outcomes of various ML methods including LR, SGD, RF, SVM, NB, KNN and DT. From the open phish website, the most effective classifier is utilized to identify dangerous websites.

In [5], they have helped by applying URL-based features in a multiclass classification context to help detect malicious URLs. [5] concentrated on three prevalent attack types such spams, dangerous malware or even phishing URLs. Their work is an additional tool in preventing such URLs. XGBoost, AdaBoost and LightGBM were the ensemble learners that were compared for performance in [5]. [5] assessed the effectiveness of a few features they called their features on various URLs. These included high-priority word-based characteristics such bag of words segmentation, Kullback-Leibler Divergence (KL Divergence), and others. They used 1,26,983 URLs from benchmark datasets to train these algorithms, and all learners gave an accuracy of 0.95 and odd. [5] gathered spam, phishing, malware, and benign URLs from publicly accessible sources. From the processed DMOZ3 dataset on the Harvard data verse website, [5] gathered 58,132 innocuous URLs. From PhishTank, 14,374 phishing URLs were gathered. 22,626 spam URLs were also gathered by [5] using publicly accessible datasets. All three of the other learners were surpassed by the AdaBoost classifier. The accuracy rates for all categorization tasks were generally quite high. When leveraging our features, XGBoost, CatBoost, and AdaBoost displayed higher accuracy rates. For all classifiers, both with and without their characteristics, the total detection accuracy for benign URLs is greater than 0.980.

Based on the dataset on in-person credit transactions, the work in [6] was completed. Vactors of catrgorical and binary data are used to represent transactions which identify the cardholder, the terminal of the transaction and the transaction itself. The card holders age, gender and the corresponding bank are taken as identifuers. The definition of the transactions are given by the amount, the mode of payment (Contactless or the PIN) and some other information that is personal to the user. The nation of the terminal and a merchant category code are taken as identifiers. If the learner is able to recognize what day the transaction was done in the set of testing data tehe transactions from both the days are taken into account and are hence very easy to seperate. An RF classifier is trained using 20,500 transactions for every day in each pair of days. The classifier is then tasked to categorize the 5000 transactions in the test set. The classification is hence evaluated, using the MCC. The covariate shift throughout th dataset's 92 days is represented by a 92x92 distance matrix which is produced by using this MCC value as a starting point. They show, with the help of the RF classifier that the precision recall of the AUC is marginally improved by 2.5% including the previously detected day type.

The uneven nature of the dataset used in [7] has a detrimental effect on the effectiveness on the models using in the ML domain. Unlike the SMOTE technique, which is used in a wide range to tackle the classes that are unbalanced, under sampling techniques like the ENN make the dataset balanced with majority of the class samples being eliminated. The proposed credit card fraud detection model generates a balanced dataset using the edited nearest neighbor (SMOTE-ENN) method and synthetic minority oversampling strategy. The hybrid resampling method known as SMOTE-ENN oversamples and under samples the data. Through the use of SMOTE and ENN, samples from the minority class are oversampled and overlapping instances are removed. This approach employs the ENN's neighborhood cleaning criterion to exclude cases in the three nearest neighbors that are different from two. LSTM models are then trained using the new balanced and resampled dataset and are then integrated with the AdaBoost technique, creating a powerful ensemble. The results from the earlier mentioned method is then combined using the "weighted voting technique" and hence producing the final results.

[8] Discusses about how domain Impersonation is employed a lot by attackers to fool victims and is done by changing minute details such as a letter etc. from the domain name such as 'google' with 'goog l'. There are more ways of deception. One-hot, a vector approach uses a vector containing the amount of characters to represent those in the URL. All of the other components of teh vector are 0 except for those that correspond to the URL. A major challenge for one-hot is that the data’s dimensionality is enormous and hence making characters very dense. The displayable characters are embeded into an s x m floating matrix where s is the number of displayable characters and m is the dimension of the vector that is embeded. As a result, the embeded vector of letters or words is closer to each other. The URL is transformed into a 2D tensor with this matrix

The suggested model in [9] uses machine learning technologies to extract and analyze numerous features of autentic and phishing URLs to detect the ones classified as phishing URLs. The algorithms that are used to identify these websites are DT, RF and SVM. [9] aims to identify such malicious URLs and narrow down to the best algorithm by analyzing the accuracy, and real and false positives.[9] employs four machine learning methods for training and extracts 12 different types of information from the structural feature of the URL of the phishing websites. . In this method, a machine learning-based website URL phishing detection system that is quicker and more accurate than prior approaches has been developed. From www.alexa.com and www.phishtank.com, respectively, URLs of trustworthy websites were gathered. Lexical features, WHOIS-based features, PageRank, Alexa Rank, and Phish tank-based features are the features that were extracted from the URL. Using Random Forest and a content-based algorithm, features

are categorized. Accuracy is used to evaluate the System's erformance. [9] uses machine learning technology to improve detection measures for phishing websites. They used a random forest algorithm, which had the lowest false positive rate, to reach 97.14% detection accuracy.

Probabilistic Neural networks, LR and Genetic programming are the three models for categorixation used in [10]. A model for unsupervised fraud transactions is mapped using SOM for credit cards. SOM model is such, that no previous knowledge is needed as it updates itself constantly with new transactions. The method has 4 parts. 2 real-world datasets are considered for evaluating proposed approach and determination of its generalizability. Fraud Transactions account for 0.172% of all transactions and come under the positive class. It is very critical to select the most important and pertinent features for detection of frauds involving credit cards. IG method with LightGBM, which helps in the decrease of dimensionality of training data is used to choose the most needed attributes. To prioritize the most crucial aspects, according to fraud and honest transactions, a Bayesian based hyperparameter is used to find the similarities between transactions. A well-integrated optimization method is used to tune the LightGBM algorithm's parameters. The fast LightGBM method can efficiently manage massive data volumes and distributed data processing. In this instance, cross validation is utilized to train and test the model in each subset of the two data sets in order to get more accurate estimates. The average of all metrics that were highlighted is determined over the whole dataset.

Summarized Information of all Literature Surveys can be found in Table 1 and Information of Accuracy, Precision and Recall in Table 2

III. PROPOSED SYSTEM

We shall be considering a dataset of about 6,51,191 URLs out of which 4,28,103 URLs are Safe or Benign, 96,457 are Defacement, 94,111 are Phishing and 32,520 are Malware. We start off by plotting a word cloud for each type of URL. Phishing URLs mimic the actual URL and contain certain keywords like HTML, HTM, and ORG. Malware URLs install unwanted software and viruses like Trojan Horse which harm the user's personal device and potentially cause a huge loss of data. Malware URLs contain certain keywords such as B4, E7, and Mozi. Defacement URLs modify the original URLs so as to make the user believe it is the actual URL. They contain keywords such as index, PHP, option, and id. Benign URLs are safe URLs that contain keywords such as Wikipedia, and YouTube. We then proceed with Feature Engineering by considering about twenty-two features such as if the URL has an IP Address if it has abnormal keywords and if the URL is indexed on Google or not which is one of the most important features which we will be looking into. The other features include counting of dots, www, @, and directory (/). embeddings (//), shortening services, HTTPS, HTTP, %, -. =, URL length, and the Host Name Length. We also consider suspicious words, digit count, letter count, and top-level directory.

All the considered features are distributed on a graph so as to give a clear picture of the distribution to the user.3 models are trained based on these features mainly Random Forest, XGBoost, and LightGBM. The accuracy of each of these machine learning models is recorded to compare which gives us the most efficient classification. The entire process involves 2 main stages, they are the training stage and the Detection Stage. The Training stage involves feature labeling by plotting the respective word clouds. The 3 models are trained based on the features that have been mentioned above. The Detection stage mainly consists of the output given by our trained model based on the features fed into it. It classifies the URLs into 4 types: Phishing, Malware, Defacement, and Benign URLs.

The fraudulent credit card transaction detection system would require us to work on unbalanced datasets, The execution process consists of 6 phases namely:

1) Collecting Data
2) Data Pre-processing
3) Data analysis
4) Separate it into training and test data
5) We feed training data to the logistic regression model and compare it with the testing data
6) Final step is to evaluate

As mentioned earlier the dataset used for fraudulent credit card transaction is unbalanced, every parameter needs to be considered carefully, and since the credit card data is sensitive a method needs to be devised to detect transactions that are fraudulent without breaching the user's privacy The data sets include the following characteristics:

• Time: - The time the transaction was made

• V1, V2, V3, V4, and V28: relates to the characteristics and limits of the credit playing cards that have been transformed into numbers.

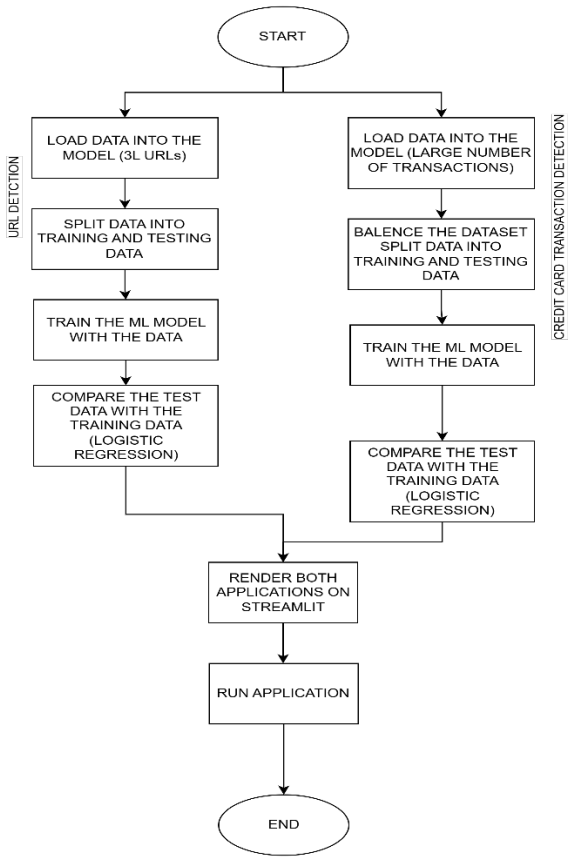

Figure 1: Shows the flow of the proposed application.

978-1-6654-9781-7/22/$31.00 ©2022 IEEE

Initially the machine learning model is trained over a dataset and tested with to attain a certain accuracy score, the accuracy of the model depends on the size of the dataset and the number of parameters or features considered for building the model. the model is later converted into a pickle file. In the front-end the necessary features are extracted from the user input and the pickle file operates on this feature to determine the result.

The entire application will be run on streamlit, an open-source framework to rapidly build Machine Learning based web applications.

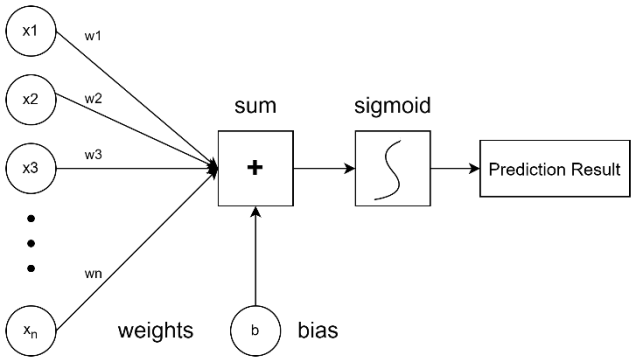

Figure 2: Shows the procedure of Logistic Regression

A. Abbreviations and Acronyms

TFIDF - Term frequency Inverse Document Frequency

RF - Random Forest

CNN - Convolution Neural Networks

DT - Decision Tree

ELM- Extreme Learning Machine

SGD-Stochastic Gradient Descent

KNN - K Nearest Neighbour

MLP - Multilayer Perceptron

SVM - Support vector Machines

SOAP - Simple Object Access Protocol

REST - Representational State Transfer

ET - Extra Tree

LR - Logistic regression

IV. CONCLUSION

All in all, we work on detecting any URL or any Transaction that is deemed to be fraud, which may cause problems to users with respect to their personal information and/or damage the resources they use. These two methodologies will be integrated to a web hosted API which can be available to anyone, having an internet connection to use and determine which transaction and which URL is fraudulent and which is benign or safe. In the modern world where development and innovations are happening rapidly everyday with technology playing a central role in this revolution, It makes technology vulnerable to malicious attacks by cyber-criminals exploiting various loopholes and vulnerabilities found in the devices and the people using these devices. Online transactions are at an all- time high, this leaves an open ground for digital thefts.

We aim at providing security against Malicious URL and against fraud credit card transactions. As we all know, safety is the basic concern for every individual and hence it is important that the user is safe while browsing on the internet.

REFERENCES

[1] Alsaedi, M., Ghaleb, F.A., Saeed, F., Ahmad, J. and Alasli, M., 2022. Cyber Threat Intelligence-Based Malicious URL Detection Model Using Ensemble Learning. Sensors, 22(9), p.3373.

[2] Prusti, D. and Rath, S.K., 2019, October. Web service based credit card fraud detection by applying machine learning techniques. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 492-497). IEEE.

[3] Ileberi, E., Sun, Y. and Wang, Z., 2021. Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost. IEEE Access, 9, pp.165286-165294.

[4] Janet, B. and Kumar, R.J.A., 2021, March. Malicious URL detection: a comparative study. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1147-1151). IEEE.

[5] Manyumwa, T., Chapita, P.F., Wu, H. and Ji, S., 2020, December. Towards Fighting Cybercrime: Malicious URL Attack Type Detection using Multiclass Classification. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 1813-1822). IEEE.

[6] Lucas, Y., Portier, P.E., Laporte, L., Calabretto, S., He-Guelton, L., Oblé, F. and Granitzer, M., 2019, June. Dataset shift quantification for credit card fraud detection. In 2019 IEEE second international conference on artificial intelligence and knowledge engineering (AIKE) (pp. 97-100). IEEE.

[7] Esenogho, E., Mienye, I.D., Swart, T.G., Aruleba, K. and Obaido, G., 2022. A neural network ensemble with feature engineering for Improved Credit Card Fraud Detection. IEEE Access, 10, pp.16400-16407.

[8] Yuan, J., Chen, G., Tian, S. and Pei, X., 2021. Malicious URL detection based on a parallel neural joint model. IEEE Access, 9, pp.9464-9472

[9] Vara, K.D., Dimble, V.S., Yadav, M.M. and Thorat, A.A., 2022. Based on URL Feature Extraction Identify Malicious Website Using Machine Learning Techniques. International Research Journal of Innovations in Engineering and Technology, 6(3), p.144.

[10] Taha, A.A. and Malebary, S.J., 2020. An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. IEEE Access, 8, pp.25579- 25587.

978-1-6654-9781-7/22/$31.00 ©2022 IEEE

Paper Number	Methodology	Limitation
[1]	RF, URL, Google, and Whois-based ANN classifier.	• CTI extracted from google • Probable Wrong information
[2]	Decision tree, Extreme learning machine, K-Nearest neighbor, Multilayer Perceptron and Support Vector Machine to find the accuracy for the fraud detection	• Only gives Accuracy of 82.58% • Can use more classification Algorithms to get better accuracy
[3]	SVM, RF and ET, XGBoost, AdaBoost etc. were used to classify benign and fraudulent credit card transactions.	• Need testing on Additional Credit card Datasets for wider range of results.
[4]	LR, SGD, RF, SVM, NB, KNN and DT algorithms used.	• Only uses selected features of a URL like Suffix, parameter length etc. to classify URLS.
[5]	Detection of Malicious URLs using XGBoost, AdaBoost, LightGBM and CatBoost Classification Algorithms.	• Need to extract more discriminative features differentiating spam
[6]	RF classifier is trained and test data is given to classify. The Matthews Correlation Coefficient is given to assess the classification.	• Only one classification algorithm is used which just gives an increase of 2.5% precision.
[7]	SMOTE-ENN sampling technique used to produce balanced dataset and worked on using Long Short-Term Memory (LSTM) technique.	• More resampling techniques and feature selection techniques for enhanced performance.
[8]	To boost the predictive effectiveness of the suggested strategy, it is necessary to use an effective parameter optimization strategy.	• Need of multiple classifications are high which requires more research and studies. • Newer, better performing versions can be selected to replace current components
[9]	DT, RF and SVM algorithms are used to detect phishing websites.	• Exploring different phishing methods that rely on lexical, network, and content-based aspects is highly necessary..
[10]	Discusses the probabilistic neural network, logistic regression, and genetic programming as three categorization techniques.	• To boost the predictive effectiveness of the suggested strategy, it is necessary to use an effective parameter optimization strategy.

Table 1: Summary of all Literature Surveys quoted in this paper

978-1-6654-9781-7/22/$31.00 ©2022 IEEE

Paper Number	Accuracy	Precision	Recall
[1]	**-**	**Using RF – 96.7% Using DT – 95.69% Using CNN – 94.4**	**-**
[2]	**Observed Accuracy – 82.58%**	**-**	**-**
[3]	**-**	**DT without SMOTE – 79.83 DT with SMOTE – 98.79%**	**DT without SMOTE – 75.75% DT with SMOTE – 99%**
[4]	**-**	**-**	**-**
[5]	**Overall Accuracy > 95%**		
[6]	**-**	**Precision-Recall saw an increase of over 2.5%**	
[7]	**-**	**-**	**-**
[8]	**-**	**-**	**-**
[9]	**Detection Accuracy - 97.14%**	**-**	**-**
[10]	**-**	**-**	**-**

Table 2: Summary of Accuracy, Precision and Recall from the papers mentioned in Literature Survey

978-1-6654-9781-7/22/$31.00 ©2022 IEEE